Big Data Education and Training

Michelle C. Dunn, Ph.D.

Senior Advisor for Data Science Training, Diversity, and Outreach

National Institutes of Health

Overarching Goal

To improve the national landscape for Big Data education and training to fulfill increasing demand for both deep analytical talent and analytical capacity for the broader workforce

What are the analytical knowledge and skills needed for Big Data?

- What is the scope of a data science curriculum?
  - Core: Informatics, computer science, statistics, mathematics, and the domain
  - Additional: Human-computer interaction/usability/design, social and behavioral science, ethics/legal

- How does it vary by audience?
  - Deep analytical vs. broader workforce
  - By domain

How can analytical training be increased for all learners?

- Traditional training
  - "Data literacy" as part of general education core requirements
  - Increased throughput for specialists

- Short-term practical experiences
  - Summer programs, internships, challenges

- Technology-based modalities
  - Online resources: online textbooks, modules, MOOCs, interactive learning environments, curricula, challenges/datasets/tools
  - Open educational resources
  - Finding training materials

Who is welcome to participate?

- Everyone.

- Learners
  - Including women, underrepresented minorities
  - cross-trainers

- Facilitators
  - Academia, government, private industry, nonprofits
  - partnerships

NIH Big Data to Knowledge

Advancing Health and Discovery through Big Data

- Training is one of 4 focus areas, launched in 2014

- Already solicits applications that address many of the modalities to increase analytical training